Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-04 09:48:42
REFERENCE NO: 213

Author Names & Affiliations

- John Moore - Internet2

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Advance Cyberinfrastructure

Title of Submission

Community Planning and Investment Coordination in Support of Science and Engineering Research

Abstract (maximum ~200 words).

Internet2 is responding to NSF's Request for Information on Future Needs for Advanced Cyberinfrastructure to Support Science and Engineering Research in order to help inform a long-term vision and strategy to advance scientific investments and accommodate technological challenges for the future. Our recommendations include: the need for greater community coordination and a defined ecosystem strategy related to services, operations, and investments; the value of an experimentation platform to inform a coordinated strategy to facilitate joint development with community member participants, align infrastructure innovation with industry leaders, lower infrastructure costs via disaggregation and open architectures, and provide a more compelling platform to support infrastructure research; and provide an environment for hands-on learning to develop the skilled workforce required by the nationwide scientific community.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Academic research in the United States relies on infrastructure developed and maintained by members of the Internet2 community - campus IT organizations, regional networks, and the Internet2 organization itself. The mission of this community is very broad, encompassing the entire spectrum of teaching, learning, and discovery that characterizes modern academia. Over the last 20 years, this infrastructure has evolved to serve the needs not only of the researchers who drive scientific discovery, but also the general needs of the academic enterprise that supports them. As this community adapts to the rapid changes of the modern world driven by technological innovation and societal evolution, it is becoming increasingly important to find ways for the various component organizations that comprise the community to work together in a more integrated fashion to optimize and improve the platform upon which the research community depends. Because this community's support mission is very broad, it is necessary to provide infrastructure and services that are generally useful and provide a clear value proposition. In addition, it is equally important to target key constituencies that require specialized services,

and strive to provide innovative and high-performance infrastructure that serves those needs and advances the national goal of advancing discovery and innovation through collaborative science.

One guidepost for supporting the latter goal is to engage with and help support key large-scale science collaboration efforts that the National Science Foundation has chosen to fund. Examples of this are the Big Data Regional Innovation Hubs, the Global Environment for Network Innovation (and the recent TIPOFF solicitation), and others. Internet2 member institutions participate directly in these programs as well as provide infrastructure that supports them. It is vital that those in the community who are providing infrastructure required to make those kinds of projects successful are engaged in understanding the requirements and are adapting and evolving the infrastructure and services they provide to serve those needs. That level of integrated planning does not generally exist today. The community provides generalized infrastructure that supports enterprise and research needs through a variety of funding mechanisms via individual budgetary processes at campuses, regional networks, and Internet2, and has managed to sustain that infrastructure over the last few decades, despite periodic funding challenges. There is some coordination and planning among these players, mainly via community members participating in Internet2 governance and advisory groups, other national organizations such as the Quilt, and regional collaborations such as the CIC.

On the research “user” side, NSF funds basic research and advanced cyberinfrastructure for specific purposes and with limited terms. In many cases, NSF requires Principal Investigators to submit sustainability plans for advanced CI, but in most cases those plans are not coordinated in any significant way with the other, more generalized community infrastructure investments. A more integrated approach would help foster the goal of creating a sustainable national cyberinfrastructure platform.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Internet2 has been engaged with the nationwide R&E community in an effort to plan for the next investment in national-scale infrastructure since early 2016. At this writing, the effort is still ongoing. The conversation is addressing the needs of the academic enterprise, as well as those of the research community. The target planning horizon is two years and beyond.

One important guiding principle that has emerged from the conversation so far has been the interest of the community in having the ability to experiment with promising ideas and technologies, and to pilot them with the goal of translating the best solutions into the production platform as quickly as possible. Although requirements are still under development, it is clear that there will be a need for a platform that is flexible and allows programmability and customization throughout the stack. In order to fully exploit such a platform, community technologists will need an appropriate framework to allow them to organize projects and use a common, reliable pilot-to-production development process.

The community has also expressed a strong desire to plan for, develop, and operationalize services at the ecosystem level, rather than relying on the current practice of limiting service delivery to organizational boundaries. These ecosystem-focused services hold the promise of delivering better end-to-end service to the research user, but require coordination between community service providers beyond what we see today.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

A. Planning and investment coordination
Planning and development of comprehensive CI solutions need to consider the entire R&E ecosystem that includes campus considerations, state and regional networks, Internet2 national-footprint infrastructure, exchange points, and international infrastructure partners. Although infrastructure investments will still be made by individual institutions, plans should be aligned among these organizations in order to make

end-to-end delivery and operational support of services consistent and reliable.

This coordinated planning and integration requires leadership to convene and drive the effort. NSF could play a key role, as can Internet2 and organizations such as the Quilt. It will be important to have general agreement as to which organizations should assume convener and coordinator roles, and how those efforts should be resourced in order to drive cooperative development among the participants in the United States, and between other international players. NSF should be part of that planning and development process and should coordinate its investments in advanced cyberinfrastructure with the base community investments. To the extent possible, NSF also should coordinate its investments with partner funding agencies in other countries so that the eventual impact and value can be global, further enhancing the potential and existence of remote collaborations and large-scale data exchange.

B. Organization and process
An agile, community-driven development process can also be significantly aided by strong R&D partnerships with commercial vendors that see the value of working closely with the community on developing new technology solutions. One can envision a spectrum where promising ideas that start as academic research into distributed systems and networking go though concept and validation phases and emerge as pilots on open, vendor supported platforms that in turn support the production needs of data-intensive science. A further step in that continuum would be for the community to step up and assume a more significant role in integrating and maintaining the various components in an open hardware and software-based platform by partnering not with individual commercial vendors but with well-managed open technology initiatives such as the Open Compute Project or the Telecom Infra Project.

C. Workforce development
As the open hardware and software environment described above becomes more prevalent, the community will need a workforce with different skill sets than exist today in the R&E community. Engineers and programmers that support advanced CI will need to be able to integrate, test, and support systems that are assembled from components, rather than relying on monolithic products that have been subject to rigorous unit, system, and field testing, and are accompanied by tech support teams and developers that can process bug fixes and feature requests.

The R&E community has a significant advantage, in that academia can produce these highly skilled individuals if there is a focused, pragmatic effort to train and mentor candidates on the infrastructure used to support learning and skills development.

Consent Statement

- “I hereby agree to give the National Science Foundation (NSF) the right to use this information for the purposes stated above and to display it on a publically available website, consistent with the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode).”